University of Zurich[UZH]

Zurich Open Repository and Archive
University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2011

The factorial structure and external validity of the prospective and retrospective memory questionnaire in older adults

Zimprich, Daniel ; Kliegel, Matthias ; Rast, Philippe

Abstract: The factorial structure of the Prospective and Retrospective Memory Questionnaire (PRMQ) was investigated in a sample of 336 older adults (aged 66-81years). Confirmatory factor analyses showed that a bifactor model of two correlated factors of prospective and retrospective memory problems and two uncorrelated group factors of positively and negatively worded items had the best fit. Such a model can be seen as a multitrait-multi-method model that separates the substantive and methodological components among the items of the PRMQ. Correlations of the four factors with external criteria (affect, neuroticism, prospective, and retrospective memory performance) revealed that the item wording factors mainly correlate with the affect variables, whereas the prospective and retrospective memory problem factors were differentially associated with memory performance. As a conceptual conclusion, these differential correlations give support to the discriminant validity of subjective prospective versus retrospective memory problems

DOI: https://doi.org/10.1007/s10433-011-0174-8

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: https://doi.org/10.5167/uzh-156742
Journal Article
Published Version

Originally published at:
Zimprich, Daniel; Kliegel, Matthias; Rast, Philippe (2011). The factorial structure and external validity of the prospective and retrospective memory questionnaire in older adults. European Journal of Ageing, 8(1):39-48.
DOI: https://doi.org/10.1007/s10433-011-0174-8

Eur J Ageing (2011) 8:39–48
DOI 10.1007/s10433-011-0174-8

ORIGINAL INVESTIGATION

The factorial structure and external validity of the prospective and retrospective memory questionnaire in older adults

Daniel Zimprich · Matthias Kliegel · Philippe Rast

Published online: 3 February 2011
© Springer-Verlag 2011

Abstract The factorial structure of the Prospective and Retrospective Memory Questionnaire (PRMQ) was investigated in a sample of 336 older adults (aged 66–81 years). Confirmatory factor analyses showed that a bifactor model of two correlated factors of prospective and retrospective memory problems and two uncorrelated group factors of positively and negatively worded items had the best fit. Such a model can be seen as a multitrait-multi-method model that separates the substantive and methodological components among the items of the PRMQ. Correlations of the four factors with external criteria (affect, neuroticism, prospective, and retrospective memory performance) revealed that the item wording factors mainly correlate with the affect variables, whereas the prospective and retrospective memory problem factors were differentially associated with memory performance. As a conceptual conclusion, these differential correlations give support to the discriminant validity of subjective prospective versus retrospective memory problems.

Keywords Prospective and retrospective memory questionnaire · Item wording · Confirmatory factor analysis

Communicated by H. W. Wahl.

D. Zimprich (✉)
Department of Gerontopsychology, Institute of Psychology, University of Zurich, Binzmühlestrasse, 14/Box 24, 8050 Zürich, Switzerland
e-mail: d.zimprich@psychologie.uzh.ch

M. Kliegel
University of Dresden, Dresden, Germany

P. Rast
University of Victoria, Victoria, BC, Canada

Prospective memory is referred to as the ability to remember delayed intentions on one's own initiative (Einstein and McDaniel 1996). Prospective memory is a common challenge of everyday life (for an overview, see Kliegel et al. 2008b). Our lives are replete with prospective memory demands, such as remembering to return a book to the library, to take medications, or to remove a pan from the burner after the food has been cooked. Conceptually, prospective memory is contrasted with the traditional topic of memory research, retrospective memory (*remembering information from the past such as facts or what we have done*; see Einstein and McDaniel 1990). As prospective memory has been suggested to be a multi-component process comprising both traditional memory elements and executive control, two components have been differentiated in prospective memory: Becoming aware *that* one had an intention in mind at the appropriate moment is referred to as the *prospective component* of prospective memory and is associated with executive attention control such as target monitoring. Retrieving the content of the specific intended action from long-term memory after becoming aware that something has to be done is referred to as the *retrospective component* of prospective memory and is associated with episodic memory networks (*what* was the intended action? Einstein and McDaniel 1996; for the discriminant validity, see Salthouse et al. 2004; Zeintl et al. 2007).

In line with its relevance for independent everyday functioning, prospective memory has increasingly begun to attract the attention of gerontology and clinical psychology, revealing substantial effects of age and several clinical conditions (e.g., autism, ADHD, Schizophrenia, Parkinson, Alzheimer's dementia) on prospective memory performance measured in the laboratory. It appears that older persons and those suffering from clinical conditions show poorer prospective memory performance (for an overview, see Kliegel et al. 2008a). So far, research has mainly focused on objective task performance with a clear emphasis on controlled laboratory tasks. Only recently, researchers have started to examine *subjective* prospective memory complaints in everyday life using standardized questionnaires. The interest in subjective prospective memory complaints in everyday life stems from the fact that if they were valid, i.e., if they would correctly reflect prospective memory problems, they could be used as early indicators of a clinical condition.

The most prominent example of such self-report measures is the *Prospective and Retrospective Memory Questionnaire* (PRMQ; Crawford et al. 2006; Smith et al. 2000; for applications in gerontology or clinical settings see, e.g., Zeintl et al. 2006). The PRMQ was designed to explicitly disentangle self-rated prospective and retrospective memory performance in everyday life and yields good reliability (Chronbach's alpha of .89; Crawford et al. 2003). Specifically, the 16-item PRMQ assesses how often errors in everyday memory tasks occur. The memory tasks described fall into different categorizations (see Table 1): (a) prospective versus retrospective, i.e., whether the memory task described is prospective or retrospective in nature, (b) short-term versus long-term, i.e., whether the memory task described taps short-term memory or long-term memory, and (c) self-cued versus environmentally cued, i.e., whether the memory task leading to a problem was triggered by oneself or by the environment. Initial psychometric analyses revealed that a three factor structure model fitted the PRMQ scores best, comprising a general memory problems factor and two orthogonal group factors of PM and RM problems (see Crawford et al. 2003).

For the present study, we extended the categorisations of the item contents of the PRMQ by including both positively and negatively worded items. In previous studies, only negatively worded items in form of questions have been administered that focused on memory failures and had a loss-orientation (e.g., "Do you forget to buy something you planned to buy, like a birthday card, even when you see the shop?" or "Do you fail to mention or give something to a visitor that you were asked to pass on?"; Crawford et al. 2003; Rönnlund et al. 2008).[1]) It is well known, however, that, in general, self-reports may be affected by several types of response sets (Cronbach 1946) and, in particular, that in aging, a loss-oriented perspective may influence self-reports and even performance (Cavanaugh et al. 1998; Hess et al. 2003). Some of the more prominent response sets discussed in psychological research are, for example, social desirability, acquiescence, evasiveness, extremity, or deviation (Cohen et al. 1996). In an attempt to prevent acquiescence and other method effects, many researchers construct questionnaires with both positively and negatively worded items. Marsh (1996, p. 817) argued in favor of equally proportioned questionnaires with respect to positively and negatively worded items. He reasoned that "without this balance, it is difficult to establish how much of the distinction between different factors is due to differences in the underlying constructs being measured as opposed to method effects." In fact, some researchers have asserted that using logically opposite items is a necessary procedure in scale development (e.g., Paulhus 1991). As will be demonstrated below, researchers can incorporate method effects into confirmatory factor analyses of measurement models, thus taking their influence into account explicitly (e.g., Zimprich et al. 2005).

Moreover, because explicitly modeled method effects can be linked to other variables, the construct validity of both the measure in question (PRMQ) and the method effects (positively and negatively worded items) can be examined. Whereas for some investigators method effects merely reflect "noise", others consider them to be meaningful (Motl and DiStefano 2002). In what follows, both the substantive factors in the PRMQ and method effects in the PRMQ will be correlated to external criterion variables, namely, actual memory performance in prospective and retrospective memory tests, and—as it has been shown that subjective memory ratings are strongly related to affect and personality (e.g., Zeintl et al. 2006; Kliegel and Zimprich 2005; Zimprich et al. 2003)—also to positive and negative affect and neuroticism. More specifically, we expected memory problems to be negatively associated with (prospective and retrospective) memory performance, reflecting that there is a certain amount of "truth" in memory complaints. Regarding affect and neuroticism, two different correlational patterns with memory problems and method effects may be distinguished. In accordance with previous studies, memory problems may be correlated with affect and neuroticism. In contrast, affect and neuroticism may be correlated with method effects, implying that method effects capture an affective component in the way items are answered.

An additional issue that will be considered in the present study is that the items of the PRMQ require participants to answer on a 5-point ordered-categorical Likert-type scale response format. This is a common feature of questionnaires, most of which consisting of ordered-categorical items. However, the assignment of integers to categories is often driven by convenience and convention rather than by a formal model. Oftentimes, integers are arbitrary and one might have little confidence that the additive relations necessary for interval-level scales are faithfully represented. Thus, for example, the integer relation $2–1 = 5–4$ does not necessarily imply that empirically *disagree–*

[1] The original items can be retrieved from the PRMQ website at http://www.psy.ed.ac.uk/research/hcn/PRMQclick/document_view.

Table 1 Categorisations of the 16 PRMQ items

Item no.	Prospective vs. retrospective	Short- vs. long-term	Self-cued vs. envir. cued	Positive vs. negative wording
1	Prospective	Short-term	Self-cued	Positive
2	Retrospective	Long-term	Envir. cued	Positive
3	Prospective	Short-term	Envir. cued	Negative
4	Retrospective	Short-term	Self-cued	Negative
5	Prospective	Long-term	Self-cued	Negative
6	Retrospective	Short-term	Envir. cued	Negative
7	Prospective	Long-term	Envir. cued	Negative
8	Retrospective	Long-term	Self-cued	Negative
9	Retrospective	Long-term	Envir. cued	Negative
10	Prospective	Short-term	Envir. cued	Positive
11	Retrospective	Short-term	Self-cued	Positive
12	Prospective	Long-term	Envir. cued	Positive
13	Retrospective	Short-term	Envir. cued	Positive
14	Prospective	Long-term	Self-cued	Positive
15	Retrospective	Long-term	Self-cued	Positive
16	Prospective	Short-term	Self-cued	Negative

Note: "envir. cued" = environmentally cued. All items were answered on a 5-point Likert type scale ranging from 1 = "agree strongly" to 5 = "disagree strongly"

strongly disagree = *strongly agree–agree*. Hence, the difference equality, which holds among the numbers that were assigned to categories and which implies measurements on an interval scale, does not necessarily hold among the verbal categories and the way people interpret them. If such items are factor-analyzed as if they were continuous or interval-scaled, there may be a critical mismatch between the information represented by the numbers assigned to the Likert-type scales and the nature of the factor model parameters on which statistical tests are based (Shadish et al. 2005). Besides the limitations arising from a levels-of-measurement perspective, another potential problem associated with ordered-categorical variables is that frequently they show departures from both univariate and multivariate normality. Previous studies have shown that this typically results in considerable negative bias of parameters and standard errors (DiStefano 2002). Although this problem appears to be less severe if item parcels or sum scores are used, once *individual items* are factor-analyzed, there are benefits in treating item-level Likert data as ordered-categorical (Muthén and Kaplan 1985). Factor analysis models for ordered-categorical variables date back to the seminal work of Christofferson (1975), who described an approach for dichotomous variables. Subsequently, Bartholomew (1980) and Muthén (1984), among others, considered the more general case of ordered-categorical variables with two or more categories. In the present study, factor analysis of ordered-categorical variables as implemented in Mplus (Muthén and Muthén 2004) was employed.

The aims of the present study were two-fold. First, the factorial structure of the PRMQ was investigated. Specifically, a sequence of increasingly complex factor models was estimated. This sequence of models started with a general factor model, followed by four different two-factor models, each accounting for one of the four different categorisations of the PRMQ items (see Table 1). Here, the goal was to examine which of the four categorizations best captured the data. Subsequently, like in Crawford et al. (2003), bifactor models were tested, that is, a general factor of memory problems was combined with four different factor pairs, each accounting for one of the four different categorisations of the PRMQ items (see Table 1). Here, the goal was to investigate which of the four categorizations was the most important once memory problems were accounted for. Eventually, a final model of prospective and retrospective memory problems combined with effects of positive and negative wordings was estimated, which fit the data best. The second aim of our study was to correlate the factors obtained from the previous analyses with external criteria. Here, the goal was to shed some light on the validity and the usefulness of the accepted factorial model of the PRMQ.

Materials and methods

Participants

Complete PRMQ data were gathered from $N = 336$ older persons participating at the second measurement occasion

(September 2006) of the Zurich Longitudinal Study on Cognitive Aging (ZULU; Zimprich et al. 2008). On average, participants were 74.4 years old (SD 4.4, 66–81 years). 153 participants (45.6%) were female. Typically, participants had 13 years of formal education. For the participants, there were no signs of cognitive impairments (as assessed by the Mini Mental Status Examination; Folstein et al. 1975) or pronounced depressive affect (as measured by the Geriatric Depression Scale; Yesavage et al. 1983). Subjective health was mostly judged as "good" or "very good" and, in addition, participants did not report any severe hearing or vision impairment. Of the original ZULU sample at T1 (March 2005), 28 participants had not returned for the second measurement occasions. Selectivity analyses showed that those who did not return at T2 differed from those who stayed in ZULU only with respect to verbal learning, but effect sizes were small (around 2% of explained variance, see Zimprich and Rast 2009). Due to the small attrition rate (8%), the sample of those who returned at T2 can be considered as being as representative as the original sample.

Measures

Prospective and retrospective memory problems

Part of the data collection protocol of ZULU at the second measurement occasion was a German version of the PRMQ (Smith et al. 2000). The PRMQ consists of 16 items, each tapping a specific memory problem, e.g., "Do you fail to mention or give something to a visitor that you were asked to pass on?" Compared to previous studies of the PRMQ, in the present study a fourth categorisation was introduced by using both positively and negatively worded items. Half of the 16 items were positively worded whereas the other half was negatively worded by turning the questions of the original PRMQ into positive and negative statements (see Table 1). For example, Item 1 of the original PRMQ ("Do you decide to do something in a few minutes' time and then forget to do it?") was reformulated into the positively worded statement "If I decide to do something in a few minutes' time, I seldom forget to do it", which was required to be answered on a five-point Likert scale ranging from 1 = "agree strongly" to 5 = "disagree strongly." The exact wordings of the positively worded items are given in the "Appendix". Positively worded items were reversed prior to the analysis such that higher scores reflected more pronounced prospective and retrospective memory problems.

External criterion variables

Previous studies have shown that subjective memory problems are associated with affect variables and neuroticism (Kliegel and Zimprich 2005; Zeintl et al. 2006; Zimprich et al. 2003). Thus, in addition to self-reported memory problems, positive and negative affect were assessed using the PANAS (Watson and Clark 1988). Neuroticism was measured using the 12 according items from the NEO-FFI (Costa and McCrae 1992).

Objective memory test performance was assessed using standard procedures. Specifically, retrospective memory performance was examined using a paired associates task consisting of 12 semantically unrelated word pairs (for details, see Zimprich et al. 2008). Prospective memory performance was measured using a Red Pencil task, a Token task, and a Background Color Changing task (for details, see Zeintl et al. 2006).

Statistical modeling

Different factor models of the PRMQ were examined using confirmatory factor analysis. In total, 10 models were examined, each reflecting different assumptions regarding the factorial structure of the PRMQ. Model 1 was a single factor model, where each item loaded on one common factor. Model 2a was a model of two correlated factors, where the first factor was defined by those items tapping prospective memory problems whereas the second factor was defined by the retrospective memory items (see Table 1, 2nd column). Model 2b was a model of two correlated factors, where the first factor was specified by short-term memory problem items whereas the second factor contained the long-term memory problem items (Table 1, 3rd column). Model 2c also was a model of two correlated factors, where the first factor was defined by items referring to self-cued memory problems and the second factor comprised those items referring to environmentally cued problems (Table 1, 4th column). Finally, Model 2d was a model of two correlated factors where the positively worded items loaded on one first factor and the negatively worded items loaded on the other factor (Table 1, 5th column).

Models 2a–d were extended by turning them into bifactor models (Yung et al. 1999). Briefly, in a bifactor representation, each item is allowed to have a positive loading on a general trait that is assumed to underlie all the items. Typically, this general factor will be conceptually broader, and is the trait that the researcher is most interested in scaling individuals on. In addition, each item can load on one "group" factor. These group factors will tend to be more conceptually narrow. In most applications, a bifactor model is specified so that the general factor and the group factors are all orthogonal to each other, i.e., they are uncorrelated. Accordingly, for Models 3a to 3b, Models 2a to 2d were extended. In Model 3a there was a general factor, on which all 16 items loaded, and two orthogonal group factors, one comprising the prospective and one comprising the retrospective items (see Crawford et al. 2003). Analogously, in Model 3b one general factor and two uncorrelated group factors of short-term and long-term memory problems were specified. Model 3c included a general factor and two orthogonal group factors of self-cued versus environmentally cued memory problem items. Finally, Model 3d consisted of a general factor and two uncorrelated group factors of items with positive and negative wordings. A final model, Model 4, represented a combination of Models 2a and 2d, combining a prospective and a retrospective factor with two uncorrelated group factors of positively and negatively worded items.

All analyses were conducted using Mplus (Muthén and Muthén 2004). Due to the items of the PRMQ being in a Likert scale format, we used confirmatory factor analyses for ordered-categorical variables (Bartholomew 1980) and the mean-corrected Weighted Least Squares Estimator (WLSM; see Flora and Curran 2004). The absolute goodness-of-fit of models was evaluated using the χ^2-test and two additional criteria, the Comparative Fit Index (CFI) and the Standardized Root Mean Square Residual (SRMR). The SRMR is the standardized difference between the observed and the predicted covariances. Values of the CFI above .95 are considered to be adequate, whereas for the SRMR values should be less than .05 for an acceptable model (cf. Hu and Bentler 1999). In comparing relative fit, we used the χ^2-difference test with the Satorra–Bentler correction in order to account for possible non-normality of the ordered-categorical data.

Table 2 Fit indices of the confirmatory factor analysis models ($N = 336$)

	χ^2_{S-B}	df	$\Delta\chi^2_{S-B}$	Δdf	SC	CFI	SRMR
Model 1	691.67*	104			0.643	0.901	0.084
Model 2a	604.42*	103	41.54*[a]	1[a]	0.635	0.915	0.078
Model 2b	687.97*	103	3.13[a]*	1[a]	0.644	0.901	0.084
Model 2c	691.35*	103	0.32[a]*	1[a]	0.643	0.901	0.084
Model 2d	589.61*	103	147.47*[a]	1[a]	0.645	0.918	0.077
Model 3a	358.52*	88	189.95*[b]	15[b]	0.588	0.954	0.067
Model 3b	473.71*	88	171.91*[c]	15[c]	0.593	0.935	0.069
Model 3c	520.67*	88	148.38*[d]	15[d]	0.602	0.927	0.071
Model 3d	270.24*	88	202.61*[e]	15[e]	0.563	0.969	0.049
Model 4	231.41*	87	17.78*[f]	1[f]	0.554	0.976	0.046

Note: df = degrees of freedom, χ^2_{S-B} = Satorra–Bentler corrected chi-square, $\Delta\chi^2_{S-B}$ = Satorra–Bentler corrected chi-square difference, *SC* Scaling Correction Factor, *CFI* Comparative Fit Index, *SRMR* Standardized Root Mean Square Residual. Model 1 is a single-factor model. Models 2 all have two correlated factors. In 2a the factors are prospective and retrospective memory, in 2b the factors refer to short-term versus long-term memory problems, 2c distinguishes between self-cued versus environmentally cued items, and 2d included a factor of positively and a factor of negatively worded items. Models 3 added a general factor to the two-factor models with the corresponding alphabetical letter—all factors were uncorrelated. Model 4 combined Models 2a and 2d

* $p < .05$

[a] Represents the difference to Model 1

[b] Represents the difference to Model 2a

[c] Represents the difference to Model 2b

[d] Represents the difference to Model 2c

[e] Represents the difference to Model 2d

[f] Represents the difference to Model 3d

Results

Confirmatory factor analyses started with Model 1, the single-factor model. As can be seen from Table 2, Model 1 did not achieve an adequate fit as judged by both the CFI and the SRMR. The average standardized loading on the single factor was 0.52. The 16 items shared about 27% variance. The standardized factor loading was particularly low for Item 16 (0.11), which means that this item ("I often forget to tell someone something I had meant to mention a few minutes ago") was relatively distinct from the remaining 16 items of the PRMQ.[2]

In Model 2a, two correlated factors of prospective versus retrospective memory problems were specified (for the categorisation of items, see Table 1). Doing so increased the absolute fit, although according to the CFI and the SRMR fit was still not adequate (see Table 2). Compared to the single-factor model (Model 1), fit had improved considerably, implying that a two-factor structure of retrospective and prospective captured the associations among the 16 items much better than a single factor. The average standardized loading of the prospective items was 0.58, and the eight items referring to prospective memory problems shared about 37% of common variance. Again, it was Item 16 which had the lowest standardized loading (0.11), implying that it was virtually unrelated to the other items. In contrast, the average standardized loading on the retrospective factor was 0.53. The items referring to retrospective memory problems had about 30% of variance in common. The correlation between both factors was .81 (or 66% shared variance), which shows that prospective and retrospective memory problems had a strong tendency to co-occur in older adults.

[2] For reasons of comparison, Model 1 was re-estimated while treating the PRMQ items as interval-scaled. Fit indices were lower (CFI = 0.773, SRMR = 0.106), as was the average standardized factor loading (.43). One reason for this may be that most items exhibited considerable skewness, which in case of only a few answer categories makes the ordered-categorical approach more suitable (cf. Muthén and Kaplan 1985).

In Model 2b, a two-factor model of items referring to short-term versus long-term memory problems was estimated (see Table 1 for the categorisation). As shown in Table 2, Model 2b did not have an adequate fit according to the CFI and the SRMR. Compared to Model 1, the improvement in fit was not significant, implying that Model 2b did not better represent the data. The correlation between the two factors was .97, showing that participants did not differentiate between short-term and long-term memory problems.

The next model, Model 2c, comprised two factors of items containing self-cued versus environmentally cued memory problems (see Table 1). Table 2 shows that the CFI and the SRMR did not index an acceptable fit of Model 2c. Compared to Model 1, there was practically no change in model fit, implying that Model 2c did not capture the associations among the 16 items better than Model 1 did. The self-cued factor and the environmentally cued factor almost collapsed with a correlation of .99, which clearly demonstrates that for the older participants the distinction between different types of cues in conjunction with memory problems was virtually irrelevant.

Subsequently, Model 2d was estimated, which included a factor of positively worded items and a correlated factor of negatively worded items (for the categorisation of items, see Table 1). Again, fit was not yet acceptable according to both the CFI and the SRMR (see Table 2). At the same time, Model 2d represented an improvement compared to Model 1. The average standardized loading of the positively worded items was 0.48. The items shared about 26% of variance. Once more, it was Item 16 which exhibited the lowest loading (0.13). In contrast, the negatively worded items had, on average, a standardized loading of 0.64. The negative-worded items shared about 42% of variance. The two factors correlated .78, implying that positively and negatively worded items tended to be answered according to their content, but not exclusively so: The fact that the correlation between the two factors was not perfect necessarily implies that the type of wording (positive vs. negative) had an influence on the way an item was answered.

The first bifactor model, Model 3a, supplemented Model 2a by a general factor. At the same time, all three factors—the general factor and the two group factors of prospective and retrospective memory items—were required to be orthogonal, i.e., uncorrelated. Table 2 shows that Model 3a exhibited an adequate fit as judged by the CFI, but not so according to the SRMR. Compared to the reference Model 2a, fit had improved considerably. The average standardized loading on the general factor was 0.51, ranging from 0.18 (Item 16) to 0.81 (Item 8). The average standardized loading of the eight prospective items on their group factor was 0.31, ranging from −0.16 (Item 2) to 0.54 (Item 7). Thus, Item 2 had a negative loading on the prospective memory problems group factor. In contrast, the average standardized loading on the group factor of retrospective memory problems was only 0.04. Four items had positive loadings (Items 2, 11, 13, 15), while the other four items had negative loadings (Items 4, 6, 8, 9). Note that the pattern of positive and negative loadings on the retrospective memory problems factor corresponded perfectly to the categorisation of the items according to their wordings.

Next, Model 3b was estimated, which combined two group factors of short-term versus long-term memory problems and a general factor in a bifactor model. According to Table 2, Model 3b did not achieve an acceptable fit, although it represented a considerable improvement compared to its reference model, Model 2b. Similarly, Model 3c with two group factors of self-cued versus environmentally cued memory problems and a general factor did not fit acceptably, though compared to its reference, Model 2c, it fitted significantly better (see Table 2). These results imply that even in a bifactor type of model the distinction between short-term versus long-term memory problems or between self-cued versus environmentally cued memory problems did not adequately describe the PRMQ data in older adults. It was irrelevant to the participants.

Finally, Model 3d was specified, which combined a general factor with two group factors of positive and negative item wordings. Model 3d had an acceptable fit both with respect to the CFI and the SRMR (Table 2). Compared to the reference model, Model 2d, the fit of Model 3d represented a statistically significant improvement of fit. The average standardized loading on the general factor was 0.49, ranging from 0.03 (Item 16) to 0.76 (Item 10). The average standardized loading of the positively worded items on their group factor was 0.35, ranging from 0.05 (Item 14) to 0.61 (Item 13). In contrast, the average standardized loading on the group factor of negatively worded items was 0.18, ranging from –0.15 (Item 7) to 0.49 (Item 6). Upon inspection, it turned out that the negatively worded items referring to prospective memory problems (Items 3, 5, 7, 16) had the smallest negative loadings on the group factor of negative wordings.

In a final model, Model 4, Models 2a and 2d were combined by specifying two correlated general factors, one of prospective and one of retrospective memory problem items, and two uncorrelated group factors, one of positively and one of negatively worded items. As Table 2 shows, Model 4 achieved an acceptable fit, which, moreover, represented a statistically significant improvement compared to Model 3d. The average standardized loading on the prospective memory problems factor was 0.56, ranging from 0.05 (Item 16) to 0.76 (Item 12). On the retrospective memory problems factor, the average standardized loading

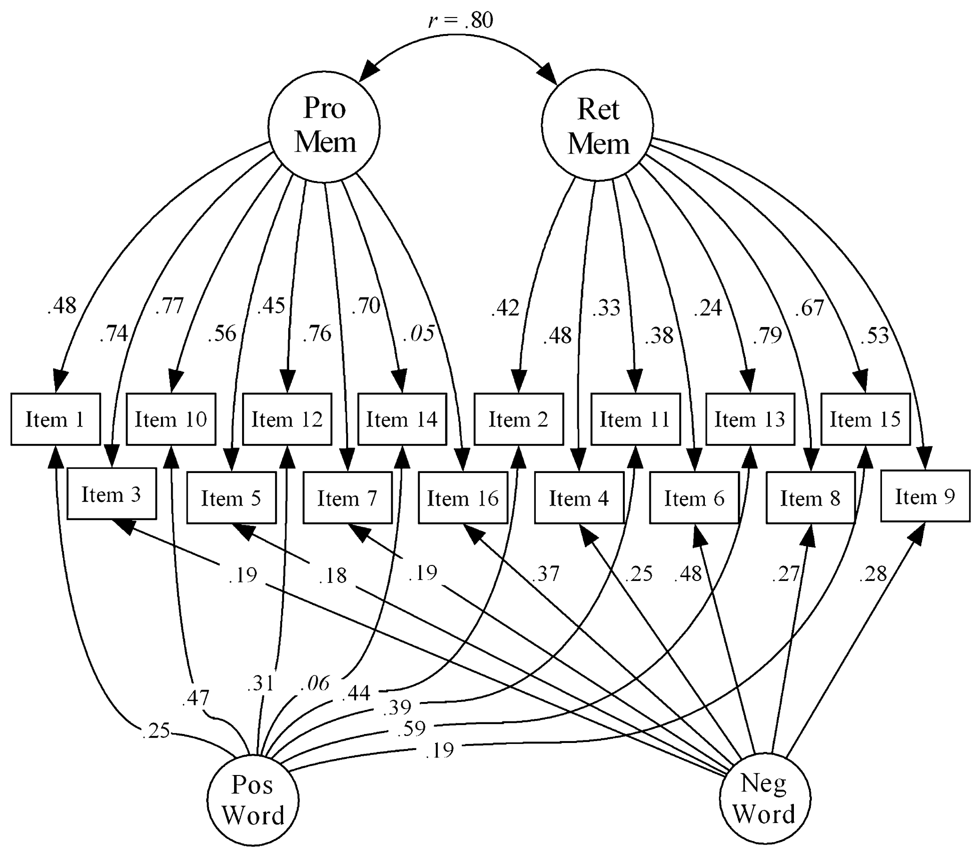

Fig. 1 Accepted Model of the PRMQ (Model 4). *Pro Mem* factor of prospective memory problems, *Ret Mem* factor of retrospective memory problems, *Pos Word* factor of positively worded items, *Neg Word* factor of negatively worded items. All parameters are standardized. Parameters in italics are not statistically significant at $p < .05$

was 0.48, ranging from 0.24 (Item 13) to 0.79 (Item 8). The prospective and retrospective memory problems factors correlated .80. For the two group factors, an average standardized loading of 0.34 emerged for the positively worded items, whereas the negatively worded items had an average standardized loading of 0.28. Due to its adequate fit and its interpretability, we decided to accept Model 4 as the final model. Model 4 is graphically represented in Fig. 1.

In order to shed some light on the validity of the PRMQ, the four factors of Model 4 were correlated with positive and negative affect, neuroticism, paired associates performance, prospective memory performance, and age. Results are shown in Table 3. The factor of self-reported prospective memory problems correlated negatively with paired associates performance and prospective memory performance. The factor of retrospective memory problems correlated negatively with paired associates performance. The tendency to answer positively worded items in a non-affirmative manner (implying more memory problems) was negatively correlated with positive affect and positively correlated with negative affect, neuroticism, and age. The tendency to answer negatively worded items in an affirmative manner (implying more memory problems) was negatively correlated with positive affect. It was positively correlated with neuroticism and age. According to Cohen's (1988) standard, effect sizes of these associations were in the medium range.

Discussion

The present study aimed at examining the factorial structure of the PRMQ (Smith et al. 2000) in a sample of 336 older adults by employing factor analysis of ordered-categorical data (Bartholomew 1980; Muthén 1984). Moreover, the PRMQ items were reformulated as positively or negatively worded items in order to shed some light on possible method effects that are so common in self-report data (cf. Cohen et al. 1996). Results show that a model of two substantive factors, prospective and retrospective memory problems, and two method factors, positive and negative wording, fit the data best. These results differ from previous findings in two respects. First, compared to Crawford et al. (2003), there was no general factor of memory problems. Rather, two distinct, albeit strongly correlated factors of prospective and retrospective memory problems emerged. Note, however, that despite of approximately 64% of shared variance, the two factors tap different aspects of memory problems. Second, extending the Crawford et al. (2003) model of the PRMQ, two method factors described the tendency to endorse positively

Table 3 Correlations of the four factors with external criteria

	Factors			
	Pro Mem	Ret Mem	Pos Word	Neg Word
Positive affect	*–.08*	*.05*	–.18	–.20
Negative affect	*.11*	*.13*	.21	*.15*
Neuroticism	*.13*	*.15*	.23	.32
Paired associates	–.24	–.28	*.06*	*–.09*
Prospective memory	–.32	*–.13*	*–.15*	*–.06*
Age	*.09*	*.07*	.17	.29

Note: *Pro Mem* factor of prospective memory problems, *Ret Mem* factor of retrospective memory problems, *Pos Word* factor of positively worded items, *Neg Word* factor of negatively worded items. Correlations in italics are not statistically significant at $p < .05$

worded items and to reject negatively worded items—independent of prospective and retrospective memory problems. Because originally the PRMQ consists of negatively worded items only, it might be that the relationships among items that have the same response format have been inflated, which may have led to a general memory problem factor. In order to verify this interpretation, though, a direct comparison of different response formats is necessary.

Before the implications of these results are discussed, a note regarding the factor analysis method used in the present study seems in order. A distinctive feature of the present research was that we applied factor analysis of ordered-categorical data. Given that Likert-type items are not necessarily interval-scaled and often exhibit strong departures from normality, instead of applying "standard" factor analysis, factor analysis of ordered-categorical variables appears more adequate. It relies on the probit model in assuming that a latent response variate that is normally distributed underlies the ordered-categorical variables (cf. Bartholomew 1980; Muthén 1984). In that sense, the probit model and the factor analysis based on it nicely fit into the assumption that Likert-type scale items—although observed responses are discrete realizations of only a small number of categories—measure a theoretically continuous construct. Although one could treat Likert-type scale items as if they were continuous, previous research has demonstrated that this typically leads to parameters and standard errors being negatively biased, especially if the number of categories is small and the distribution shows departures from normality (DiStefano 2002; Muthén and Kaplan 1985). Although the relevant statistical methodology of factor analysis of ordered-categorical data has been developed three decades ago (Christofferson 1975), it does not seem to have had a broader impact on factor analysis practice when it comes to Likert-scaled items. With the implementation of the appropriate methodology in major software packages (e.g., Muthén and Muthén 2004), however, factor analysis of ordered-categorical variables is now becoming more available as a data-analytic tool (cf. Flora and Curran 2004; Rast et al. 2009). In a broader sense, in the present study we illustrated the capabilities of the factor analysis model of ordered-categorical variables as an analytical framework for linking both theoretical and methodological considerations in examining a self-reported measure and method effects of item wording.

Turning to the effects of item wordings, the present study has also shown that once the items of the PRMQ were formulated positively and negatively, the resulting method effect of wording needed to be included in the measurement model to achieve adequate fit. This finding is of importance in several respects: First, from Model 2d one might conclude that two method factors described the data better than the other three categorisations. That is, the method effect of item wordings preponderated the categorisation of short-term and long-term memory problems and that of environmentally cued versus self-cued problems. Regarding the distinction between prospective and retrospective memory problems, the same picture emerged as long as only two factors were allowed in the models. In the final model (Model 4), however, the combination of two correlated substantive factors of prospective and retrospective memory problems with two group factors of positive and negative item wordings fitted best. The strongest loading on the Positive Wording factor was that of Item 13 ("I seldom look at something without realising that I have seen it moments before"), while Item 6 ("I often fail to recognise a character in a radio or television show from scene to scene") had the strongest loading on the Negative Wording factor. Both method factors were negatively correlated with positive affect, while there was a positive association with neuroticism. This result implies that irrespective of the level of self-reported memory problems those persons being more neurotic and feeling less positive affect were more likely to reject positively worded items and to endorse negatively worded items. Such a finding is in line with Watson and Pennebaker's (1989) surmise that a variety of personality traits exist that are all related to symptom complaints because they might be considered indicators of neuroticism. Both method factors were positively related to age, which means that independent of their memory problems older persons tended to endorse negatively worded items and reject positively worded items. Although the reported correlations appear modest by traditional standards, they fall in the middle of those found in personality research (Hemphill 2003) and those reported for age changes in personality (Allemand et al. 2008). Moreover, taking into account the comparatively small age range and the fact that method factors represent relatively narrow group factors, one might consider the correlations with affect, neuroticism, and age even remarkable. At the same time, these results show that method effects do not merely reflect "noise." To the contrary, due to their association with affect variables and age in the present study, if they remain uncontrolled for they might inflate the relations of the conceptually broader general factors of prospective and retrospective memory problems with age and affect, which are frequently found in research on self-reported memory problems (Kliegel and Zimprich 2005; Zeintl et al. 2006; Zimprich et al. 2003).

It remains unclear why Item 16 did hardly show any relations with the remaining items, which led to low and even non-significant factor loadings. In previous studies (Crawford et al. 2003), Item 16 did not exhibit such an unusual pattern. Conceptually closest to Item 16 are Item 1 (which has a positive wording) and Item 4 (which is retrospective in nature), and the correlations of Item 16 are relatively strongest with these two items. However, in terms of effect size, correlations were in the small range (.19 and .16). It might be that older adults interpreted Item 16 less of a prospective memory problem, but rather as a symptom of distractibility, such as being absentminded or easily disturbed in one's focused attention or as an interrupted sequence of cognitive actions (cf. Rast et al. 2009). Unfortunately, because Item 16 was the only item with the specific combination of categorizations, such an interpretation must remain speculative.

Together, a relatively complex four-factor model of the 16 items emerged for the PRMQ. In fact, Model 4 may be viewed as a multitrait-multimethod model (Marsh and Grayson 1995), which separates and empirically estimates the substantive and methodological[3] components among the items of the PRMQ. One consequence of taking method effects into account by using both positively and negatively worded items might be that in the present study, self-reported prospective and retrospective memory problems loaded on two different factors. In previous studies they represented group factors in the presence of a general factor of memory problems (cf. Crawford et al. 2003, 2006; Rönnlund et al. 2008). However, prospective and negative memory problems were strongly correlated in the present study, implying that both types of problems co-occurred in older adults, at least as judged by themselves. Notwithstanding, the correlations of both factors with external criteria showed that they, despite their overlap, in part, represented distinguishable memory capacities. While prospective memory problems were negatively associated with both paired associates performance (a typical retrospective memory test) *and* prospective memory performance, retrospective memory problems were selectively related to retrospective memory test performance. This nicely dovetails with a conceptual model of prospective memory as a multi-component process comprising both, a prospective and a retrospective memory component (see Einstein and McDaniel 1996). Hence, as an important conceptual conclusion, the differential correlation structure underlines previous experimental results on the discriminant validity of prospective versus retrospective memory tasks also for subjective memory processes (Salthouse et al. 2004; Zeintl et al. 2007).

[3] Model 4 was also re-estimated treating the PRMQ items as interval-scaled. Again, although findings were virtually the same, fit indices were lower (CFI = 0.935, SRMR = 0.069). Once more, this shows that treating the PRMQ items as ordered-categorical not only is statistically more suitable but also comes with the benefit of a better model fit.

Acknowledgments The Zurich Longitudinal Study on Cognitive Aging (ZULU) is supported by the Swiss National Science Foundation, grants SNSF-100013-103525 and SNSF-100014-122613/1.

Appendix

Wordings of the positively worded items used in the study

Item no.	Wording
1	If I decide to do something in a few minutes' time, I seldom forget to do it
2	I hardly fail to recognise a place I have visited before
10	If I intend to take something with me, before leaving a room or going out, I rarely leave it behind
11	I rarely mislay something that I have just put down, like a magazine or glasses
12	I seldom fail to mention or give something to a visitor that I was asked to pass on
13	I seldom look at something without realising that I have seen it moments before
14	If I tried to contact a friend or relative who was out, I hardly forget to try again later
15	I rarely forget what I watched on television the previous day

References

Allemand M, Zimprich D, Hendriks AAJ (2008) Age differences in five personality domains across the lifespan. Develop Psychol 44:758–770

Bartholomew DJ (1980) Factor analysis for categorical data. J R Stat Soc B 42:293–321

Cavanaugh JC, Feldman JM, Hertzog C (1998) Memory beliefs as social cognition: a reconceptualization of what memory questionnaires assess. Rev Gen Psychol 2:48–65

Christofferson A (1975) Factor analysis of dichotomized variables. Psychometrika 40:5–32

Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd edn. Lawrence Erlbaum, Hillsdale, NJ

Cohen RJ, Swerdlik ME, Phillips SM (1996) Psychological testing and assessment: an introduction to tests and measurement. Mayfield, Mountain View, CA

Costa PT Jr, McCrae RR (1992) Professional manual: revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI). PAR, Odessa, FL

Crawford JR, Smith G, Maylor EA, Della Sala S, Logie RH (2003) The prospective and retrospective memory questionnaire (PRMQ): normative data and latent structure in a large non-clinical sample. Memory 11:261–275

Crawford JR, Henry JD, Ward AL, Blake J (2006) The prospective and retrospective memory questionnaire (PRMQ): latent structure, normative data and discrepancy analysis. Br J Clin Psychol 45:83–104

Cronbach LJ (1946) Response sets and test design. Educ Psychol Measur 6:475–494

DiStefano C (2002) The impact of categorization with confirmatory factor analysis. Structural Equation Modeling 9:327–346

Einstein GO, McDaniel MA (1990) Normal aging and prospective memory. J Exp Psychol Learn Mem Cogn 16:717–726

Einstein GO, McDaniel MA (1996) Retrieval processes in prospective memory: theoretical approaches and some new findings. In: Brandimonte M, Einstein GO, McDaniel MA (eds) Prospective memory: theory and applications. Erlbaum, Mahwah, pp 115–142

Flora DB, Curran PJ (2004) An empirical evaluation of alternative methods of estimation for factor analysis with ordinal data. Psychol Methods 9:466–491

Folstein MF, Folstein SE, McHugh PR (1975) Mini-mental-state: a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 12:189–198

Hemphill JF (2003) Interpreting the magnitudes of correlation coefficients. Am Psychol 58:78–79

Hess TM, Auman C, Colcombe SJ, Rahhal TA (2003) The impact of stereotype threat on age differences in memory performance. J Gerontol 58B:3–11

Hu L, Bentler PM (1999) Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Model 6:1–55

Kliegel M, Zimprich D (2005) Predictors of cognitive complaints in older adults: a mixture regression approach. Eur J Aging 2:13–23

Kliegel M, Jäger T, Altgassen M, Shum D (2008a) Clinical neuropsychology of prospective memory. In: Kliegel M, McDaniel MA, Einstein GO (eds) Prospective memory: cognitive, neuroscience, developmental, and applied perspectives. Erlbaum, Mahwah, pp 283–308

Kliegel M, McDaniel MA, Einstein GO (eds) (2008b) Prospective memory: cognitive, neuroscience, developmental, and applied perspectives. Erlbaum, Mahwah

Marsh HW (1996) Positive and negative self-esteem: a substantively meaningful distinction or artifactors? J Pers Soc Psychol 70:810–819

Marsh HW, Grayson D (1995) Latent variable models of multitrait–multimethod data. In: Hoyle RH (ed) Structural equation modeling: concepts, issues, and applications. Sage, Thousand Oaks, CA, pp 177–198

Motl RW, DiStefano C (2002) Longitudinal invariance of self-esteem and method effects associated with negatively worded items. Struct Equ Model 9:562–578

Muthén BO (1984) A general structural equation model with dichotomous, ordered-categorical, and continuous latent variable indicators. Psychometrika 49:115–132

Muthén BO, Kaplan D (1985) A comparison of some methodologies for the factor-analysis of non-normal Likert variables. Br J Math Stat Psychol 38:171–180

Muthén LK, Muthén BO (2004) Mplus user's guide. Muthén & Muthén, Los Angeles, CA

Paulhus DL (1991) Measurement and control of response bias. In: Robinson JP, Shaver PR, Wrightsman LS (eds) Measures of personality and social psychological attitudes. Academic Press, San Diego, CA, pp 17–59

Rast P, Zimprich D, van Boxtel M, Jolles J (2009) Factorial structure and measurement invariance of the Cognitive Failures Questionnaire across the adult lifespan. Assessment 16:145–158

Rönnlund M, Mäntyla T, Nilsson L–G (2008) The prospective and retrospective memory questionnaire (PRMQ): factorial structure, relations to global subjective memory ratings, and Swedish norms. Scand J Psychol 49:11–18

Salthouse TA, Berish DE, Siedlecki KL (2004) Construct validity and age sensitivity of prospective memory. Mem Cogn 32:1133–1148

Shadish WR, Cook TD, Campbell DT (2005) Experimental and quasi-experimental designs for generalized causal inference. Houghton–Mifflin, Boston, MA

Smith G, Della Sala S, Logie R, Maylor EA (2000) Prospective and retrospective memory in normal aging and dementia: a questionnaire study. Memory 8:311–321

Watson D, Clark LA (1988) Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol 54:1063–1070

Watson D, Pennebaker JW (1989) Health complaints, stress, and distress: exploring the central role of negative affectivity. Psychol Rev 96:233–253

Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M et al (1983) Development and validation of a geriatric depression screening scale: a preliminary report. J Psychiatr Res 17:37–49

Yung YF, Thissen D, McLeod LD (1999) On the relationship between the higher-order factor model and the hierarchical factor model. Psychometrika 64:113–128

Zeintl M, Kliegel M, Rast P, Zimprich D (2006) Prospective memory complaints can be predicted by prospective memory performance in older adults. Dement Geriatr Cogn Disord 22:209–215

Zeintl M, Kliegel M, Hofer SM (2007) The role of processing resources in prospective and retrospective memory within old age. Psychol Aging 22:826–834

Zimprich D, Rast P (2009) Verbal learning changes in older adults across 18 months. Aging Neuropsychol Cogn 16:461–484

Zimprich D, Martin M, Kliegel M (2003) Subjective cognitive complaints, memory performance, and depressive affect in old age: a change-oriented approach. Int J Aging Hum Dev 57: 341–368

Zimprich D, Perren S, Hornung R (2005) A two-level confirmatory factor analysis of a modified Rosenberg self-esteem scale. Educ Psychol Measur 65:465–481

Zimprich D, Martin M, Kliegel M, Dellenbach M, Rast P, Zeintl M (2008) Cognitive abilities in old age: results from the Zurich longitudinal study on cognitive aging. Swiss J Psychol 67:177–195